1. Description of Methods

Rosenbrock Function:

For simple problem1, I implemented the Nesterov Momentum descent method. I initially attempted to solve the problem using Gradient Descent and backtracking line search for my parameter alpha, but I often found myself exceeding the count limit. With the realization that the Rosenbrock function has a sizeable valley floor, I began to try momentum methods (with no line-search I also save precious counts). Nesterov seemed like a good choice given its extra precaution taken to reduce oscillations and my need for achieving minimization in as few of iterations as possible (given that the count was capped to 20). Alpha and beta were found by holding one constant while tuning the other and checking performance.

Himmelblau Function:

For simple problem2, I opted to use Nesterov Momentum again. It worked effectively for Rosenbrock and appeared to be a good choice to power through the no man's land base that is sizeable in the function. Because of this larger base, it became no surprise that my beta value was larger (compared to when solving Rosenbrock), while my alpha was nearly the same.

Powell Function:

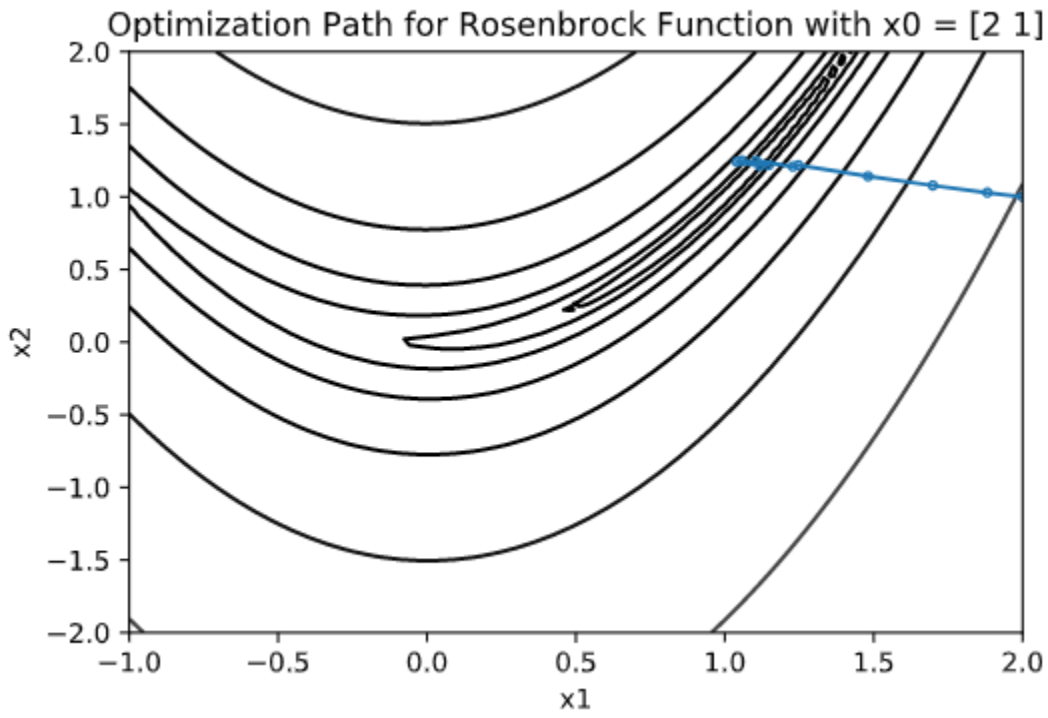
For simple problem3, I initially tried Nesterov Momentum. However, after some tuning, I still wasn't satisfied with the results and tried generic Momentum (not nesterov). Surprisingly, this gave me exemplary results. I'm not exactly sure why this occurred, but my guess is somewhere between some poor tuning with nesterov on my part or a unique quality about this function that I'm overlooking (is it especially flat somewhere?).

Secret Functions:

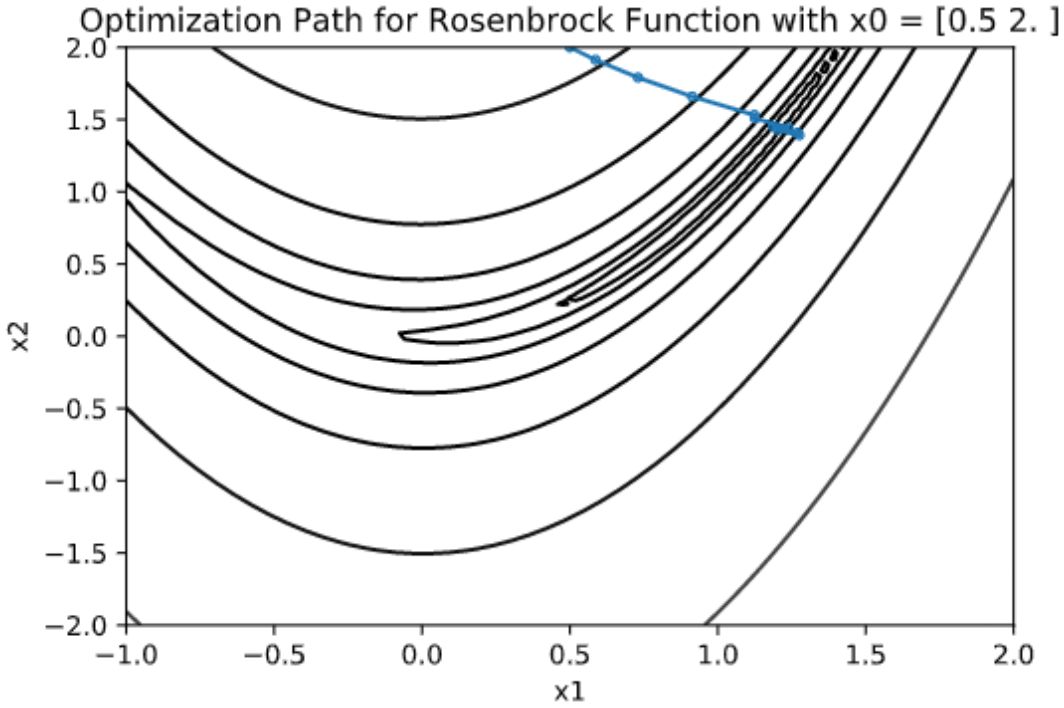
Given the great performance of Nesterov with 2/3 of my simple problems, I elected to try it for the secret functions and per the autograder, it held up pretty well!

2. Rosenbrock Contour Plots with Optimization Paths

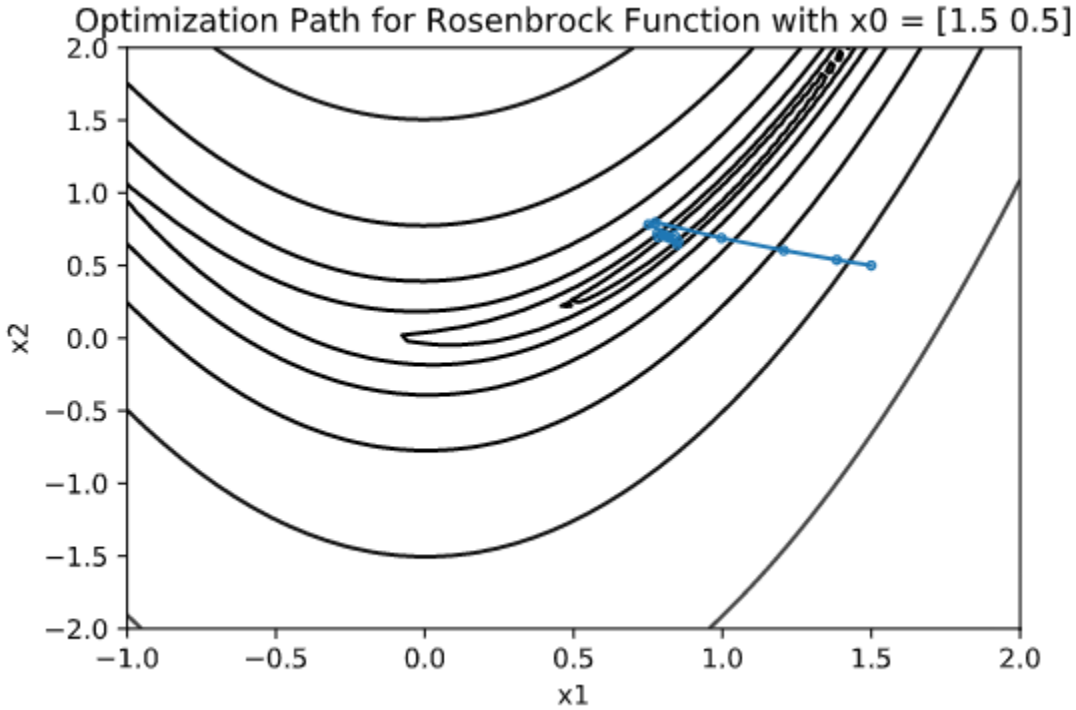
Plot 1:



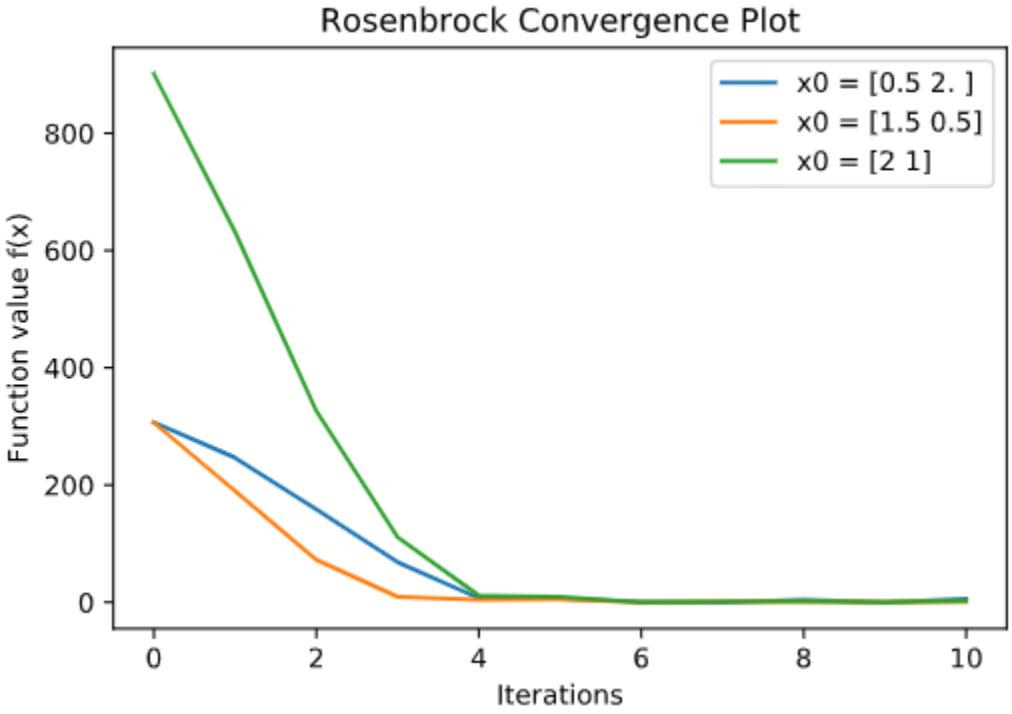
Plot 2:

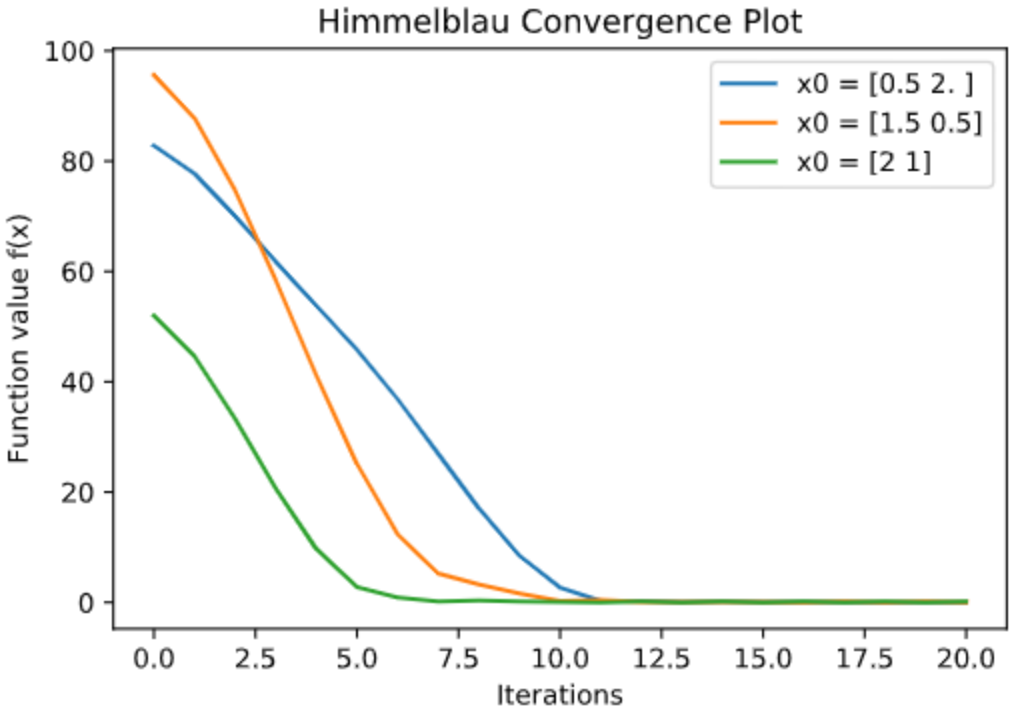

Optimization Path for Rosenbrock Function with x0 = [0.5 2.]
x2
x1
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

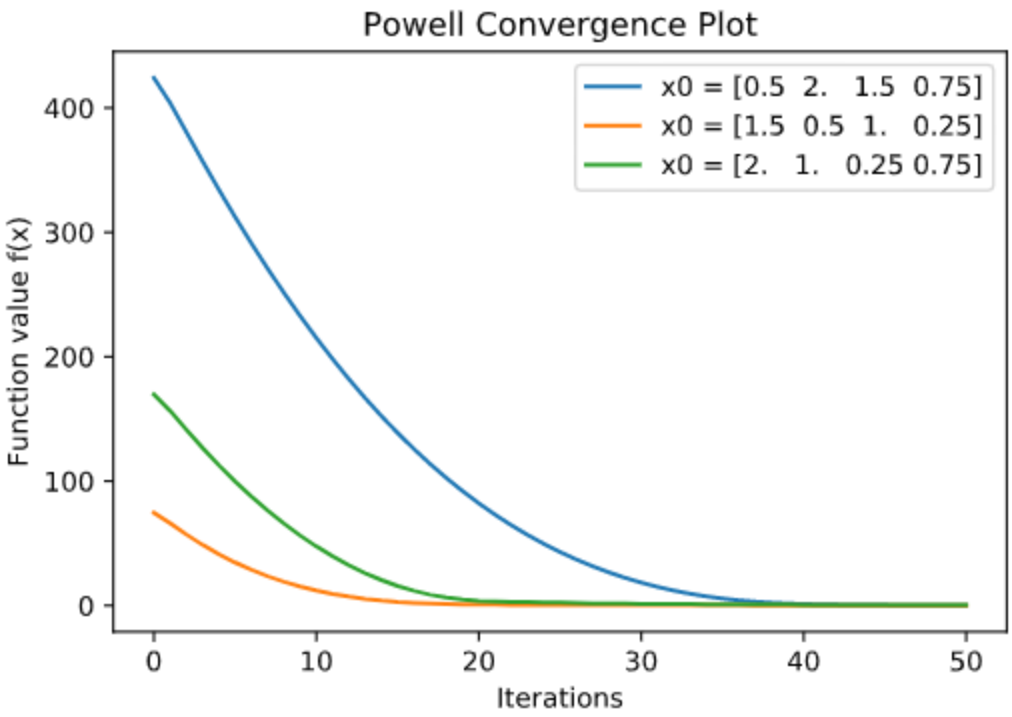
Plot 3:


Optimization Path for Rosenbrock Function with x0 = [1.5 0.5]
x2
x1
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

3. Convergence Plots






Powell Convergence Plot
x0 = [0.5 2. 1.5 0.75]
x0 = [1.5 0.5 1. 0.25]
x0 = [2. 1. 0.25 0.75]
Function value f(x)
400
300
200
100
0
0
10
20
30
40
50
Iterations